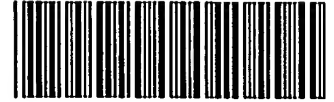


## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/522,356  
Source: P45/10  
Date Processed by STIC: 11/18/05

# ***ENTERED***



RT

## RAW SEQUENCE LISTING

DATE: 11/18/2005

PATENT APPLICATION: US/10/522,356

TIME: 10:25:17

Input Set : A:\12286155.app

Output Set: N:\CRF4\11182005\J522356.raw

3 <110> APPLICANT: WHITELAW, CHRISTOPHER BRUCE ALEXANDER  
 4 CLARK, ANTHONY JOHN  
 5 WOLF, CHARLES ROLAND  
 7 <120> TITLE OF INVENTION: MULTI-REPORTER GENE MODEL FOR TOXICOLOGICAL SCREENING  
 9 <130> FILE REFERENCE: 102286.155 US1  
 11 <140> CURRENT APPLICATION NUMBER: 10/522,356  
 12 <141> CURRENT FILING DATE: 2005-01-26  
 14 <150> PRIOR APPLICATION NUMBER: PCT/GB03/003192  
 15 <151> PRIOR FILING DATE: 2003-07-25  
 17 <150> PRIOR APPLICATION NUMBER: GB 0217402.7  
 18 <151> PRIOR FILING DATE: 2002-07-26  
 20 <160> NUMBER OF SEQ ID NOS: 41  
 22 <170> SOFTWARE: PatentIn version 3.3  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 10  
 26 <212> TYPE: PRT  
 27 <213> ORGANISM: Homo sapiens  
 29 <400> SEQUENCE: 1  
 30 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  
 31 1 5 10  
 34 <210> SEQ ID NO: 2  
 35 <211> LENGTH: 14  
 36 <212> TYPE: PRT  
 37 <213> ORGANISM: Simian virus  
 39 <400> SEQUENCE: 2  
 40 Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
 41 1 5 10  
 44 <210> SEQ ID NO: 3  
 45 <211> LENGTH: 9  
 46 <212> TYPE: PRT  
 47 <213> ORGANISM: Unknown  
 49 <220> FEATURE:  
 50 <223> OTHER INFORMATION: Description of Unknown Sequence: Haemagglutinin  
 51 epitope from unknown organism  
 53 <400> SEQUENCE: 3  
 54 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala  
 55 1 5  
 58 <210> SEQ ID NO: 4  
 59 <211> LENGTH: 12  
 60 <212> TYPE: PRT  
 61 <213> ORGANISM: Unknown  
 63 <220> FEATURE:  
 64 <223> OTHER INFORMATION: Description of Unknown Sequence: Clone100

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,356

DATE: 11/18/2005

TIME: 10:25:17

Input Set : A:\12286155.app

Output Set: N:\CRF4\11182005\J522356.raw

```

65      epitope from unknown organism
67 <400> SEQUENCE: 4
68 Asn Val Arg Phe Ser Thr Ile Val Arg Arg Arg Ala
69 1          5          10
72 <210> SEQ ID NO: 5
73 <211> LENGTH: 14
74 <212> TYPE: PRT
75 <213> ORGANISM: Unknown
77 <220> FEATURE:
78 <223> OTHER INFORMATION: Description of Unknown Sequence: Rab11a
79      epitope from unknown organism
81 <400> SEQUENCE: 5
82 Lys Gln Met Ser Asp Arg Arg Glu Asn Asp Met Ser Pro Ser
83 1          5          10
86 <210> SEQ ID NO: 6
87 <211> LENGTH: 15
88 <212> TYPE: PRT
89 <213> ORGANISM: Unknown
91 <220> FEATURE:
92 <223> OTHER INFORMATION: Description of Unknown Sequence: DOB
93      epitope from unknown organism
95 <400> SEQUENCE: 6
96 Ser Gly Asn Glu Val Ser Arg Ala Val Leu Leu Pro Gln Ser Cys
97 1          5          10          15
100 <210> SEQ ID NO: 7
101 <211> LENGTH: 17
102 <212> TYPE: PRT
103 <213> ORGANISM: Unknown
105 <220> FEATURE:
106 <223> OTHER INFORMATION: Description of Unknown Sequence: SG11
107      epitope from unknown organism
109 <400> SEQUENCE: 7
110 Ser Ser Leu Ser Tyr Thr Asn Pro Ala Val Ala Ala Thr Ser Ala Asn
111 1          5          10          15
114 Leu
118 <210> SEQ ID NO: 8
119 <211> LENGTH: 15
120 <212> TYPE: PRT
121 <213> ORGANISM: Unknown
123 <220> FEATURE:
124 <223> OTHER INFORMATION: Description of Unknown Sequence: ErbB4
125      epitope from unknown organism
127 <400> SEQUENCE: 8
128 Arg Ser Thr Leu Gln His Pro Asp Tyr Leu Gln Glu Tyr Ser Thr
129 1          5          10          15
132 <210> SEQ ID NO: 9
133 <211> LENGTH: 16
134 <212> TYPE: PRT
135 <213> ORGANISM: Unknown

```

## RAW SEQUENCE LISTING

DATE: 11/18/2005

PATENT APPLICATION: US/10/522,356

TIME: 10:25:17

Input Set : A:\12286155.app

Output Set: N:\CRF4\11182005\J522356.raw

```

137 <220> FEATURE:
138 <223> OTHER INFORMATION: Description of Unknown Sequence: ARF
139     epitope from unknown organism
141 <400> SEQUENCE: 9
142 Val Ser Thr Leu Leu Arg Trp Glu Arg Phe Pro Gly His Arg Gln Ala
143 1           5           10           15
146 <210> SEQ ID NO: 10
147 <211> LENGTH: 20
148 <212> TYPE: PRT
149 <213> ORGANISM: Unknown
151 <220> FEATURE:
152 <223> OTHER INFORMATION: Description of Unknown Sequence: RYK
153     epitope from unknown organism
155 <400> SEQUENCE: 10
156 Lys Phe Gln Gln Leu Val Gln Cys Leu Thr Glu Phe His Ala Ala Leu
157 1           5           10           15
160 Gly Ala Tyr Val
161           20
164 <210> SEQ ID NO: 11
165 <211> LENGTH: 20
166 <212> TYPE: PRT
167 <213> ORGANISM: Unknown
169 <220> FEATURE:
170 <223> OTHER INFORMATION: Description of Unknown Sequence: WILPEP1
171     epitope from unknown organism
173 <400> SEQUENCE: 11
174 Gln Glu Gln Cys Gln Glu Val Trp Arg Lys Arg Val Ile Ser Ala Phe
175 1           5           10           15
178 Leu Lys Ser Pro
179           20
182 <210> SEQ ID NO: 12
183 <211> LENGTH: 14
184 <212> TYPE: PRT
185 <213> ORGANISM: Unknown
187 <220> FEATURE:
188 <223> OTHER INFORMATION: Description of Unknown Sequence: HAF10
189     epitope from unknown organism
191 <400> SEQUENCE: 12
192 Arg Leu Ser Asp Lys Thr Gly Pro Val Ala Gln Glu Lys Ser
193 1           5           10
196 <210> SEQ ID NO: 13
197 <211> LENGTH: 543
198 <212> TYPE: DNA
199 <213> ORGANISM: Mus musculus
201 <220> FEATURE:
202 <221> NAME/KEY: CDS
203 <222> LOCATION: (1)..(540)
205 <400> SEQUENCE: 13
206 atg aag atg ctg ctg ctg ctg tgt ttg gga ctg acc cta gtc tgt gtc

```

48

## RAW SEQUENCE LISTING

DATE: 11/18/2005

PATENT APPLICATION: US/10/522,356

TIME: 10:25:17

Input Set : A:\12286155.app

Output Set: N:\CRF4\11182005\J522356.raw

```

207 Met Lys Met Leu Leu Leu Leu Cys Leu Gly Leu Thr Leu Val Cys Val
208 1 5 10 15
210 cat gca gaa gaa gct agt tct acg gga agg aac ttt aat gta gaa aag 96
211 His Ala Glu Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys
212 20 25 30
214 att aat ggg gaa tgg cat act att atc ctg gcc tct gac aaa aga gaa 144
215 Ile Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu
216 35 40 45
218 aag ata gaa gat aat ggc aac ttt aga ctt ttt ctg gag caa atc cat 192
219 Lys Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His
220 50 55 60
222 gtc ttg gag aaa tcc tta gtt ctt aaa ttc cat act gta aga gat gaa 240
223 Val Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu
224 65 70 75 80
226 gag tgc tcg gaa tta tct atg gtt gct gac aaa aca gaa aag gct ggt 288
227 Glu Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly
228 85 90 95
230 gaa tat tct gtg acg tat gat gga ttc aat aca ttt act ata cct aag 336
231 Glu Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys
232 100 105 110
234 aca gac tat gat aac ttt ctt atg gct cat ctc att aac gaa aag gat 384
235 Thr Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp
236 115 120 125
238 ggg gaa acc ttc cag ctg atg ggg ctc tat ggc cga gaa cca gat ttg 432
239 Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu
240 130 135 140
242 agt tca gac atc aag gaa agg ttt gca caa cta tgt gag aag cat gga 480
243 Ser Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys His Gly
244 145 150 155 160
246 atc ctt aga gaa aat atc att gac cta tcc aat gcc aat cgc tgc ctc 528
247 Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu
248 165 170 175
250 cag gcc cga gaa tga 543
251 Gln Ala Arg Glu
252 180
255 <210> SEQ ID NO: 14
256 <211> LENGTH: 180
257 <212> TYPE: PRT
258 <213> ORGANISM: Mus musculus
260 <400> SEQUENCE: 14
261 Met Lys Met Leu Leu Leu Leu Cys Leu Gly Leu Thr Leu Val Cys Val
262 1 5 10 15
265 His Ala Glu Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys
266 20 25 30
269 Ile Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu
270 35 40 45
273 Lys Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His
274 50 55 60
277 Val Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/522,356

DATE: 11/18/2005

TIME: 10:25:17

Input Set : A:\12286155.app

Output Set: N:\CRF4\11182005\J522356.raw

```

278 65          70          75          80
281 Glu Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly
282          85          90          95
285 Glu Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys
286          100          105          110
289 Thr Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp
290          115          120          125
293 Gly Glu Thr Phe Gln Leu Met Gly Leu Tyr Gly Arg Glu Pro Asp Leu
294          130          135          140
297 Ser Ser Asp Ile Lys Glu Arg Phe Ala Gln Leu Cys Glu Lys His Gly
298 145          150          155          160
301 Ile Leu Arg Glu Asn Ile Ile Asp Leu Ser Asn Ala Asn Arg Cys Leu
302          165          170          175
305 Gln Ala Arg Glu
306          180
309 <210> SEQ ID NO: 15
310 <211> LENGTH: 540
311 <212> TYPE: DNA
312 <213> ORGANISM: Artificial Sequence
314 <220> FEATURE:
315 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
316      recombinant mMUP reporter molecule
318 <220> FEATURE:
319 <221> NAME/KEY: CDS
320 <222> LOCATION: (1)..(537)
322 <400> SEQUENCE: 15
323 ggg ccc ctg gga tcc atg gag cag aaa ctc atc tct gaa gag gat ctg      48
324 Gly Pro Leu Gly Ser Met Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
325 1          5          10          15
327 acc atg gaa gct agt tct acg gga agg aac ttt aat gta gaa aag att      96
328 Thr Met Glu Ala Ser Ser Thr Gly Arg Asn Phe Asn Val Glu Lys Ile
329          20          25          30
331 aat ggg gaa tgg cat act att atc ctg gcc tct gac aaa aga gaa aag      144
332 Asn Gly Glu Trp His Thr Ile Ile Leu Ala Ser Asp Lys Arg Glu Lys
333          35          40          45
335 ata gaa gat aat ggc aac ttt aga ctt ttt ctg gag caa atc cat gtc      192
336 Ile Glu Asp Asn Gly Asn Phe Arg Leu Phe Leu Glu Gln Ile His Val
337          50          55          60
339 ttg gag aaa tcc tta gtt ctt aaa ttc cat act gta aga gat gaa gag      240
340 Leu Glu Lys Ser Leu Val Leu Lys Phe His Thr Val Arg Asp Glu Glu
341 65          70          75          80
343 tgc tcg gaa tta tct atg gtt gct gac aaa aca gaa aag gct ggt gaa      288
344 Cys Ser Glu Leu Ser Met Val Ala Asp Lys Thr Glu Lys Ala Gly Glu
345          85          90          95
347 tat tct gtg acg tat gat gga ttc aat aca ttt act ata cct aag aca      336
348 Tyr Ser Val Thr Tyr Asp Gly Phe Asn Thr Phe Thr Ile Pro Lys Thr
349          100          105          110
351 gac tat gat aac ttt ctt atg gct cat ctc att aac gaa aag gat ggg      384
352 Asp Tyr Asp Asn Phe Leu Met Ala His Leu Ile Asn Glu Lys Asp Gly

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/522,356

DATE: 11/18/2005  
TIME: 10:25:18

Input Set : A:\12286155.app  
Output Set: N:\CRF4\11182005\J522356.raw

Base Note:

Presence of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> <223> fields of each sequence which presents at least one n or Xaa.

Seq#:36; N Pos. 5,6,7,8,9

Seq#:40; N Pos. 6,7,8

Seq#:41; N Pos. 6,7,8

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/522,356

DATE: 11/18/2005

TIME: 10:25:18

Input Set : A:\12286155.app

Output Set: N:\CRF4\11182005\J522356.raw

1565 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36 after pos.:0  
1626 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:0  
1645 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41 after pos.:0